

**METHOD AND SYSTEM FOR PROCESSING DOWNSTREAM PACKETS OF  
AN OPTICAL NETWORK**

5

**STATEMENT REGARDING RELATED APPLICATIONS**

This application is a continuation-in-part of a non-provisional patent application entitled, "System and Method for Communicating Optical Signals between a Data Service Provider and Subscribers," filed on July 5, 2001 and assigned U.S. Application Serial No. 09/899,410. The present application is also related to non-provisional application entitled, "System and Method for Communicating Optical Signals Upstream and Downstream between a Data Service Provider and Subscribers," filed on October 4, 2001 and assigned U.S. Serial No. 09/971,363. The present application claims priority to provisional patent application entitled, "Systems to Provide Video, Voice and Data services via Fiber Optic Cable - Part 2," filed on October 26, 2000 and assigned U.S. Application Serial No. 60/244,052; provisional patent application entitled, "Systems to Provide Video, Voice and Data services via Fiber Optic Cable - Part 3," filed on December 28, 2000 and assigned U.S. Application Serial No. 60/258,837; provisional patent application entitled, "Protocol to Provide Voice and Data Services via Fiber Optic Cable," filed on October 27, 2000 and assigned U.S. Application Serial No. 60/243,978; and provisional patent application entitled, "Protocol to Provide Voice and Data Services via Fiber Optic Cable-Part 2," filed on May 8, 2001 and assigned U.S. Application Serial No. 60/289,112, the entire contents of each of these applications are also incorporated by reference.

25

**TECHNICAL FIELD**

The present invention relates to video, voice, and data communication. More particularly, the present invention relates to a system and method for communicating downstream optical signals from a data service provider to one or more subscribers.

## BACKGROUND OF INVENTION

The increasing reliance on communication networks to transmit more complex data, such as voice and video traffic, is causing a very high demand for bandwidth. To  
5 resolve this demand for bandwidth, communication networks are relying more upon optical fibers to transmit this complex data. Conventional communication architectures that employ coaxial cables are slowly being replaced with communication networks that comprise only fiber optic cables. One advantage that optical fibers have over coaxial cables is that a much greater amount of information can be carried on an optical fiber.

10 The Fiber-to-the-home (FTTH) optical network architecture has been a dream of many data service providers because of the aforementioned capacity of optical fibers that enable the delivery of any mix of high-speed services to businesses and consumers over highly reliable networks. Related to FTTH is fiber to the business (FTTB). FTTH and FTTB architectures are desirable because of improved signal quality, lower maintenance,  
15 and longer life of the hardware involved with such systems. However, in the past, the cost of FTTH and FTTB architectures have been considered prohibitive. But now, because of the high demand for bandwidth and the current research and development of improved optical networks, FTTH and FTTB have become a reality.

A conventional hybrid fiber-to-the-home (FTTH)/hybrid fiber-coax (HFC)  
20 architecture has been proposed by the industry. HFC is currently the architecture of choice for many cable television systems. In this FTTH/HFC architecture, an active signal source is placed between the data service hub and the subscriber. Typically, in this architecture, the active source comprises a router. This conventional router typically has multiple data ports that are designed to support individual subscribers. More specifically,  
25 the conventional router uses a single port for each respective subscriber. Connected to each data port of the router is an optical fiber which, in turn, is connected to the subscriber. The connectivity between data ports and optical fibers with this conventional FTTH/HFC architecture yield a very fiber intensive last mile. It noted that the terms, "last mile" and "first mile", are both generic terms used to describe the last portion of an  
30 optical network that connects to subscribers.

In addition to a high number of optical cables originating from the router, the FTTH/HFC architecture requires radio frequency signals to be propagated along traditional coaxial cables. Because of the use of coaxial cables, numerous radio frequency (RF) amplifiers are needed between the subscriber and the data service help. For example, RF amplifiers are typically needed every one to three kilometers in a coaxial type system.

The use of coaxial cables and the FTTH/HFC architecture adds to the overall cost of the system because two separate and distinct networks are present in such an architecture. In other words, the FTTH/HFC architecture has high maintenance cost because of the completely different wave guides (coaxial cable in combination with optical fiber) in addition to the electrical and optical equipment needed to support such two distinct systems. More simply, the FTTH/HFC architecture merely combines an optical network with an electrical network with both networks running independently of one another.

One problem with the electrical network in the FTTH/HFC architecture involves cable modem technology which supports the data communications between the data service provider and the subscriber. The data service subscriber typically employs a cable modem termination system (CMTS) to originate downstream data communications that are destined to the subscriber. To receive these downstream data communications, the subscriber will typically use a cable modem that operates according to a particular protocol known in the industry as Data-Over-Cable-Service-Interface-Specification (DOCSIS). The DOCSIS protocol defines service flows, which are identifications assigned to groups of packets by the CMTS for the downstream flows based on an inspection of a number of parameters in a packet.

More specifically, a service flow is a media access control (MAC)-layer transport service that provides unique directional transport of packets either to upstream packets transmitted by the cable modem or to downstream packets transmitted by the CMTS. The identifications assigned to groups of packets in the DOCSIS protocol can include parameters such as TCP, UDP, IP, LLC, and 802.1 P/Q identifiers contained in an incoming packet.

Based on these identifications, the CMTS assigns a service flow ID (SFID) to a particular datastream. A service flow typically exists when the CMTS assigns this SFID to a datastream. The SFID serves as the principle identifier in the CMTS for the service flow. A service flow is characterized by at least an SFID and an associated direction.

- 5 One of the main drawbacks of the DOCSIS protocol for downstream data communications is that this protocol does not offer any guaranteed bandwidth. In other words, every cable modem in a particular subscriber group competes for bandwidth in both the upstream and downstream directions when a particular modem needs it. This competition between modems for bandwidths can significantly affect the quality of
- 10 service of data communications for each individual cable modem receiving downstream data communications.

- For example, subscribers that desire to use their cable modem for T1 communications require a constant bit rate and consistent arrival time of packets in order to reduce any jitter in the communications. T1 communications can include telephone
- 15 calls, video conferencing, and other similar traffic. Because each cable modem according to the DOCSIS protocol competes for bandwidth, it is possible that some cable modems will not be provided with a constant bit rate for their T1 communications. In such a scenario, the quality of T1 communications can suffer. That is, during a telephone call or a video conference the subscriber may notice either delays in communications or
- 20 truncation in conversations with the other party to the telephone call or video conference.

- DOCSIS is designed to operate over an RF modulated network, which imposes certain restrictions on the protocol. Return bandwidth is low relative to downstream bandwidth, as a result of the way spectrum is apportioned in the two directions. This causes problems with certain applications requiring more symmetrical bandwidth. These
- 25 applications include peer-to-peer file transfer, video conferencing and communications from web servers.

- Accordingly, there is a need in the art for a system and method for communicating optical signals between a data service provider and a subscriber that eliminates the use of coaxial cables and related hardware and software necessary to support the data signals
- 30 propagating along the coaxial cables. There is also a need in the art for a system and

method for communicating optical signals between a data service provider and a subscriber that can service a large number of subscribers while reducing the number of connections at the data service hub.

There is also a need in the art for a method and system for handling downstream  
5 optical communications that can police or monitor downstream bandwidths for quality of service at exit portions of the optical network. There is a further need in the art for a system and method that can allocate additional or reduce downstream bandwidths based upon one of demand or the type of service selected by one or more subscribers of an optical network. There is also a need in the art for a method and system for controlling  
10 the volume or content (or both) of downstream optical communications that are received by subscribers of an entirely optical network.

#### SUMMARY OF THE INVENTION

The present invention is generally drawn to a system and method for efficient  
15 propagation of data and broadcast signals over an optical fiber network. More specifically, the present invention is generally drawn to a method and system for handling downstream optical communications originating from a data service hub of an optical network that are transmitted to subscribers of the optical network. The term "downstream" can define a communication direction where a data service hub originates  
20 data signals that are sent downwards towards subscribers of an optical network. Conversely, the term "upstream" can define a communication direction where a subscriber originates data signals that are sent upwards towards a data service hub of an optical network.

Unlike the conventional art which polices data at the entry points of a network, the  
25 present invention can police or monitor downstream bandwidths for quality of service at exit portions of an optical network. That is, the present invention can police downstream communication traffic near the outer edges of an optical network that are physically close to the subscribers of the optical network. In this way, the network provider can control the volume or content (or both) of downstream communications that are received by  
30 subscribers of the optical network.

To control volume or content (or both) of downstream communications, the present invention employs multiple levels of evaluation for downstream communication traffic. The multiple levels of evaluation can comprise classifying downstream packets and then evaluating whether the downstream packets match certain size and rate parameters. Specifically, a plurality of classifiers can categorize or classify downstream packets, where each classifier is typically associated with a particular policer. Each policer can also be associated with a particular output buffer that has a priority relative to other output buffers.

Each policer can receive a downstream packet from one or more classifiers. The policer can evaluate the size and rate parameters of a particular downstream packet. For example, a policer can compare a downstream packet to a peak rate, a sustained rate, and a burst size that are assigned to the policer by a network administrator. The network administrator can configure the peak rate, sustained rate, and burst size monitored by each policer to track different types of downstream packets.

If a downstream packet exceeds the peak rate assigned to a policer, then the policer can discard the downstream packet. If the downstream packet exceeds the assigned sustained rate or burst size assigned to a policer, then the policer can identify this traffic as a certain type of traffic, such as "non-conforming traffic." On the other hand, if the downstream packet matches or falls within an sustained rate or burst size of a policer, then the policer can identify this traffic as a certain type of traffic, such as "conforming traffic." The policer can then assign weighted random early discard values (such as a maximum drop probability, maximum threshold, and a minimum threshold) that are unique and separate between conforming downstream traffic and non-conforming downstream traffic. Each policer can operate as a two-stage token bucket algorithm where the first stage bucket enforces the peak rates for the downstream communication traffic. The second stage of each token bucket can identify packets that exceed the burst size or the sustained rate assigned to a particular policer.

One output buffer of several output buffers can receive a packet from respective policer. Each output buffer can separately implement a weighted random early discard (WRED) algorithm to determine if packets should be admitted to a respective buffer or

dropped. Each output buffer can use the weighted random early discard value assigned to the downstream packet in the weighted random early discard algorithm.

With the WRED algorithm and classifying traffic by type, certain communication traffic can be given a higher priority over other types of traffic. For example, subscribers that use the optical network for T1 communications require a constant bit rate and consistent arrival time of packets in order to reduce any jitter. T1 communications can include telephone calls, video conferencing, and other similar traffic. To help reduce the possibility of any jitter with the T1 communications, the present invention can assign such T1 communications a higher priority relative to other types of communication traffic that do not require constant bit rates. Other communications that do not require constant bit rates and that can be assigned a lower priority can include Internet surfing, transferring files between computers, and other similar communications.

The present invention can be implemented in hardware such as application specific integrated circuits (ASICs) or field programmable gate arrays (FPGAs) or a combination thereof. However, the present invention is not limited to hardware and can comprise software.

The present invention can comprise a transceiver node that further comprises an optical tap routing device and one or more optical tap multiplexers. The optical tap routing device can determine which optical tap multiplexer is to receive a downstream electrical signal, or identify which of the plurality of optical taps originated an upstream optical signal. The optical tap routing device can format data and implement the protocol required to send and receive data from each individual subscriber connected to a respective optical tap. The optical tap routing device can further comprise an eight-port switch.

The eight-port switch can feed into one or more optical tap multiplexers. Each optical tap multiplexer can comprise one or more packet classifiers, one or more policers, and one or more output buffers.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A is a functional block diagram of the some core components of an exemplary optical network architecture according to the present invention.

Fig. 1B is a functional block diagram illustrating exemplary functionality and a location of this exemplary functionality in a network according to the present invention.

Fig. 2 is a functional block diagram illustrating an exemplary optical network architecture for the present invention.

Fig. 3 is a functional block diagram illustrating an exemplary outdoor transceiver node according to the present invention.

Fig. 4 is a functional block diagram illustrating an optical tap connected to a subscriber interface by a single optical waveguide according to one exemplary embodiment of the present invention.

Fig. 5 is a functional block diagram illustrating an exemplary optical tap routing device coupled to an exemplary optical tap multiplexer according to the present invention.

Fig. 6 is a logic flow diagram illustrating an exemplary method for processing downstream packets leaving or exiting a network according to one exemplary embodiment of the present invention.

Fig. 7 is a logic flow diagram illustrating an exemplary sub-process for evaluating in-profile packets according to one exemplary embodiment of the present invention.

Fig. 8 is a logic flow diagram illustrating an exemplary sub-process for evaluating out-of-profile packets according to one exemplary embodiment of the present invention.

Fig. 9 is a graph illustrating weighted random early discard for out-of-profile packets according to one exemplary embodiment of the present invention.

Fig. 10 is a graph illustrating weighted random early discard for in-profile packets according to one exemplary embodiment of the present invention.



## DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

The present invention may be embodied in hardware or software or a combination thereof disposed within an optical network. The present invention can comprise a transceiver node that further comprises an optical tap routing device and a plurality of optical tap multiplexers for receiving downstream packets from the optical tap routing device. Each optical tap multiplexer may comprise a plurality of classifiers and a plurality of policers. With the classifiers and policers, the present invention can support at least one gigabit or faster data rate, and Ethernet communications in optical form to and from the data service hub and partition or apportion this optical bandwidth into distribution groups of a predetermined number. The present invention can allow optical bandwidth to be offered to subscribers in preassigned increments. The flexibility and diversity of the present invention can be attributed to a few components.

Referring now to the drawings in which like numerals represent like elements throughout the several figures, aspects of the present invention and the illustrative operating environment will be described.

Figure 1A is a functional block diagram illustrating an exemplary optical network architecture 100 according to the present invention. The exemplary optical network architecture 100 comprises a data service hub 110 that is connected to one or more outdoor transceiver nodes 120. The transceiver nodes 120, in turn, are connected to an optical taps 130. The optical taps 130 can be connected to a plurality of subscriber optical interfaces 140. Between respective components of the exemplary optical network architecture 100 are optical waveguides such as optical waveguides 150, 160, 170, and 180. The optical waveguides 150-180 are illustrated by arrows where the arrowheads of the arrows illustrate exemplary directions of data flow between respective components of the illustrative and exemplary optical network architecture 100. While only an individual transceiver node 120, an individual optical tap 130, and an individual subscriber optical interface 140 are illustrated in Figure 1A, as will become apparent from Figure 2 and its corresponding description, a plurality of transceiver nodes 120, optical taps 130, and subscriber optical interfaces 140 can be employed without departing from the scope and spirit of the present invention. Typically, in many of the exemplary embodiments of the

present invention, multiple subscriber optical interfaces 140 are connected to one or more optical taps 130.

5 The outdoor transceiver node 120 can allocate additional or reduced bandwidth based upon the demand of one or more subscribers that use the subscriber optical interfaces 140. The outdoor transceiver node 120 can be designed to withstand outdoor environmental conditions and can be designed to hang on a strand or fit in a pedestal or "hand hole." The outdoor transceiver node can operate in a temperature range between minus 40 degrees Celsius to plus 60 degrees Celsius. The transceiver node 120 can operate in this temperature range by using passive cooling devices that do not consume  
10 power.

Unlike the conventional routers disposed between the subscriber optical interface 140 and data service hub 110, the outdoor transceiver node 120 does not require active cooling and heating devices that control the temperature surrounding the transceiver node 120. The present invention attempts to place more of the decision-making electronics at  
15 the data service hub 110 instead of the transceiver node 120. Typically, the decision-making electronics are larger in size and produce more heat than the electronics placed in the transceiver node of the present invention. Because the transceiver node 120 does not require active temperature controlling devices, the transceiver node 120 lends itself to a compact electronic packaging volume that is typically smaller than the environmental enclosures of conventional routers.  
20

In one exemplary embodiment of the present invention, three trunk optical waveguides 160, 170, and 180 (that can comprise optical fibers) can conduct optical signals from the data service hub 110 to the outdoor transceiver node 120. It is noted that the term "optical waveguide" used in the present application can apply to optical fibers,  
25 planar light guide circuits, and fiber optic pigtails and other like optical waveguides.

A first optical waveguide 160 can carry broadcast video and other signals. The signals can be carried in a traditional cable television format wherein the broadcast signals are modulated onto carriers, which in turn, modulate an optical transmitter (not shown) in the data service hub 110. A second optical waveguide 170 can carry downstream targeted  
30 services such as data and telephone services to be delivered to one or more subscriber

optical interfaces 140. In addition to carrying subscriber-specific optical signals, the second optical waveguide 170 can also propagate internet protocol broadcast packets, as is understood by those skilled in the art.

5 In one exemplary embodiment, a third optical waveguide 180 can transport data signals upstream from the outdoor transceiver node 120 to the data service hub 110. The optical signals propagated along the third optical waveguide 180 can also comprise data and telephone services received from one or more subscribers. Similar to the second optical waveguide 170, the third optical waveguide 180 can also carry IP broadcast packets, as is understood by those skilled in the art.

10 The third or upstream optical waveguide 180 is illustrated with dashed lines to indicate that it is merely an option or part of one exemplary embodiment according to the present invention. In other words, the third optical waveguide 180 can be removed. In another exemplary embodiment, the second optical waveguide 170 propagates optical signals in both the upstream and downstream directions as is illustrated by the double  
15 arrows depicting the second optical waveguide 170. In such an exemplary embodiment where the second optical waveguide 170 propagates bidirectional optical signals, only two optical waveguides 160, 170 would be needed to support the optical signals propagating between the data server's hub 110 in the outdoor transceiver node 120. In another exemplary embodiment (not shown), a single optical waveguide can be the only link  
20 between the data service hub 110 and the transceiver node 120. In such a single optical waveguide embodiment, three different wavelengths can be used for the upstream and downstream signals. Alternatively, bi-directional data could be modulated on one wavelength.

In one exemplary embodiment, the optical tap 130 can comprise an 8-way optical  
25 splitter. This means that the optical tap 130 comprising an 8-way optical splitter can divide downstream optical signals eight ways to serve eight different subscriber optical interfaces 140. In the upstream direction, the optical tap 130 can combine the optical signals received from the eight subscriber optical interfaces 140.

In another exemplary embodiment, the optical tap 130 can comprise a 4-way  
30 splitter to service four subscriber optical interfaces 140. Yet in another exemplary

embodiment, the optical tap 130 can further comprise a 4-way splitter that is also a pass-through tap meaning that a portion of the optical signal received at the optical tap 130 can be extracted to serve the 4-way splitter contained therein while the remaining optical energy is propagated further downstream to another optical tap or another subscriber optical interface 140. The present invention is not limited to 4-way and 8-way optical splitters. Other optical taps having fewer or more than 4-way or 8-way splits are not beyond the scope of the present invention.

Referring now to Figure 1B, this figure illustrates exemplary functionality and a location of this exemplary functionality in a network 103 according to the present invention. The network 103 can comprise several of the components of the architecture 100 described in Figure 1A.

As noted above, unlike the conventional art which polices data at the entry points 105 of a network, the present invention can police or monitor downstream bandwidths for quality of service at exit portions 107 of an optical network 103. That is, the present invention can police downstream communication traffic near the outer edges 107 of an optical network 103 that are relatively, physically close to the subscribers (subscriber optical interfaces 140) of the optical network. In this way, the network provider can control the volume or content (or both) of downstream communications that are received by subscribers of the optical network 103.

As illustrated in Figure 1B, a third party web server 182 may be coupled to an optical network 103 that comprises transceiver nodes 120. With the transceiver nodes 120 of the present invention, the network provider can limit or control the bandwidth capacity granted to a subscriber. In other words, the network provider can control what quality of service is given to a particular subscriber (such as a subscriber optical interface 140 that may be coupled to a computer 142 running a web browser).

Specifically, the transceiver node 120 running the protocol of the present invention enables a network provider to create different tiers of service that can be ordered by the subscriber. For example, the transceiver node can offer a particular subscriber or groups of subscribers downstream bandwidth in units of 1, 2, 5, 10, 20, 50, 100, 200, and 450 Megabits per second (Mb/s) that are governed by the transceiver node

120.

Referring now to Figure 2, this Figure is a functional block diagram illustrating an exemplary optical network architecture 100 that further includes subscriber groupings 200 that correspond with a respective outdoor transceiver node 120. Figure 2 illustrates the diversity of the exemplary optical network architecture 100 where a number of optical waveguides 150 connected between the outdoor transceiver node 120 and the optical taps 130 is minimized. Figure 2 also illustrates the diversity of subscriber groupings 200 that can be achieved with the optical tap 130.

Each optical tap 130 can comprise an optical splitter. The optical tap 130 allows multiple subscriber optical interfaces 140 to be coupled to a single optical waveguide 150 that is connected to the outdoor transceiver node 120. In one exemplary embodiment, six optical fibers 150 are designed to be connected to the outdoor transceiver node 120. Through the use of the optical taps 130, sixteen subscribers can be assigned to each of the six optical fibers 150 that are connected to the outdoor transceiver node 120.

In another exemplary embodiment, twelve optical fibers 150 can be connected to the outdoor transceiver node 120 while eight subscriber optical interfaces 140 are assigned to each of the twelve optical fibers 150. Those skilled in the art will appreciate that the number of subscriber optical interfaces 140 assigned to a particular waveguide 150 that is connected between the outdoor transceiver node 120 and a subscriber optical interface 140 (by way of the optical tap 130) can be varied or changed without departing from the scope and spirit of the present invention. Further, those skilled in the art recognize that the actual number of subscriber optical interfaces 140 assigned to the particular fiber optic cable is dependent upon the amount of power available on a particular optical fiber 150.

As depicted in subscriber grouping 200, many configurations for supplying communication services to subscribers are possible. For example, while optical tap 130<sub>A</sub> can connect subscriber optical interfaces 140<sub>A1</sub> through subscriber optical interface 140<sub>AN</sub> to the outdoor laser transmitter node 120, optical tap 130<sub>A</sub> can also connect other optical taps 130 such as optical tap 130<sub>AN</sub> to the transceiver node 120. The combinations of optical taps 130 with other optical taps 130 in addition to combinations of optical taps

130 with subscriber optical interfaces 140 are limitless. With the optical taps 130, concentrations of distribution optical waveguides 150 at the transceiver node 120 can be reduced. Additionally, the total amount of fiber needed to service a subscriber grouping 200 can also be reduced.

5           With the active transceiver node 120 of the present invention, the distance between the transceiver node 120 and the data service hub 110 can comprise a range between 0 and 80 kilometers. However, the present invention is not limited to this range. Those skilled in the art will appreciate that this range can be expanded by selecting various off-the-shelf components that make up several of the devices of the present  
10       system.

          Those skilled in the art will appreciate that other configurations of the optical waveguides disposed between the data service hub 110 and outdoor transceiver node 120 are not beyond the scope of the present invention. Because of the bi-directional capability of optical waveguides, variations in the number and directional flow of the optical  
15       waveguides disposed between the data service hub 110 and the outdoor transceiver node 120 can be made without departing from the scope and spirit of the present invention.

          Those skilled in the art will appreciate that the selection of optical waveguide transceiver 430 (Figure 3) in the outdoor transceiver node 120, and the corresponding transceiver (not shown) in data service hub 110, may be optimized for the optical path  
20       lengths needed between the data service hub 110 and the outdoor transceiver node 120. Further, those skilled in the art will appreciate that the wavelengths discussed are practical but are only illustrative in nature. In some scenarios, it may be possible to use communication windows at 1310 and 1550 nm in different ways without departing from the scope and spirit of the present invention. Further, the present invention is not limited  
25       to a 1310 and 1550 nm wavelength regions. Those skilled in the art will appreciate that smaller or larger wavelengths for the optical signals are not beyond the scope and spirit of the present invention.

          Referring now to Figure 3, this Figure illustrates a functional block diagram of an exemplary outdoor transceiver node 120 of the present invention. In this exemplary  
30       embodiment, the transceiver node 120 can comprise a unidirectional optical signal input

port **405** that can receive optical signals propagated from the data service hub **110** that are propagated along a first optical waveguide **160**. The optical signals received at the unidirectional optical signal input port **405** can comprise broadcast video data. The optical signals received at the input port **405** are propagated to an amplifier **410** such as an  
5 Erbium Doped Fiber Amplifier (EDFA) in which the optical signals are amplified. The amplified optical signals are then propagated to a splitter **415** that divides the broadcast video optical signals among duplexers **420** that are designed to forward optical signals to predetermined subscriber groups **200**.

The transceiver node **120** can further comprise a bi-directional optical signal  
10 input/output port **425** that connects the transceiver node **120** to a second optical waveguide **170** that supports bi-directional data flow between the data service hub **110** and transceiver node **120**. Downstream optical signals flow through the bi-directional optical signal input/output port **425** to an optical waveguide transceiver **430** that converts downstream optical signals into the electrical domain. The optical waveguide transceiver  
15 further converts upstream electrical signals into the optical domain. The optical waveguide transceiver **430** can comprise an optical/electrical converter and an electrical/optical converter.

Downstream and upstream electrical signals are communicated between the optical waveguide transceiver **430** and an optical tap routing device **435**. The optical tap  
20 routing device **435** can manage the interface with the data service hub optical signals and can route or divide or apportion the data service hub signals according to individual tap multiplexers **440** that communicate optical signals with one or more optical taps **130** and ultimately one or more subscriber optical interfaces **140**. It is noted that tap multiplexers  
25 **440** operate in the electrical domain to modulate laser transmitters in order to generate optical signals that are assigned to groups of subscribers coupled to one or more optical taps.

Optical tap routing device **435** is notified of available upstream data packets as they arrive, by each tap multiplexer **440**. The optical tap routing device is connected to each tap multiplexer **440** to receive these upstream data packets. The optical tap routing  
30 device **435** relays the packets to the data service hub **110** via the optical waveguide

transceiver 430. The optical tap routing device 435 can build a lookup table from these upstream data packets coming to it from all tap multiplexers 440 (or ports), by reading the source IP address of each packet, and associating it with the tap multiplexer 440 through which it came. This lookup table can then be used to route packets in the downstream path. As each packet comes in from the optical waveguide transceiver 430, the optical tap routing device 435 looks at the destination IP address (which is the same as the source IP address for the upstream packets). From the lookup table the optical tap routing device 435 can determine which port is connected to that IP address, so it sends the packet to that port. This can be described as a normal layer 3 router function as is understood by those skilled in the art.

The optical tap routing device 435 can assign multiple subscribers to a single port. More specifically, the optical tap routing device 435 can service groups of subscribers with corresponding respective, single ports. The optical taps 130 coupled to respective tap multiplexers 440 can supply downstream optical signals to pre-assigned groups of subscribers who receive the downstream optical signals with the subscriber optical interfaces 140.

In other words, the optical tap routing device 435 can determine which tap multiplexer 440 is to receive a downstream electrical signal, or identify which of a plurality of optical taps 130 propagated an upstream optical signal (that is converted to an electrical signal). The optical tap routing device 435 can format data and implement the protocol required to send and receive data from each individual subscriber connected to a respective optical tap 130. The optical tap routing device 435 can comprise a computer or a hardwired apparatus that executes a program defining a protocol for communications with groups of subscribers assigned to individual ports.



The single ports of the optical tap routing device are connected to respective tap multiplexers 440. With the optical tap routing device 435, the transceiver node 120 can adjust a subscriber's bandwidth on a subscription basis or on an as-needed or demand basis. The transceiver node 120 via the optical tap routing device 435 can offer data bandwidth to subscribers in pre-assigned increments. For example, the transceiver node 120 via the optical tap routing device 435 can offer a particular subscriber or groups of subscribers bandwidth in units of 1, 2, 5, 10, 20, 50, 100, 200, and 450 Megabits per second (Mb/s). Those skilled in the art will appreciate that other subscriber bandwidth units are not beyond the scope of the present invention.

Electrical signals are communicated between the optical tap routing device 435 and respective tap multiplexers 440. The tap multiplexers 440 propagate optical signals to and from various groupings of subscribers. Each tap multiplexer 440 is connected to a respective optical transmitter 325. Each optical transmitter 325 can comprise one of a Fabry-Perot (F-P) laser, a distributed feedback laser (DFB), or a Vertical Cavity Surface Emitting Laser (VCSEL). However, other types of optical transmitters are possible and are not beyond the scope of the present invention. The optical transmitters produce the downstream optical signals that are propagated towards the subscriber optical interfaces 140.

Those skilled in the art will appreciate that the functions ascribed to the optical tap routing device 435 and the tap multiplexers 440 are exemplary in nature. In other words, functions may be performed differently than what is described. Some of the functions performed by the routing device 435 could be performed by the tap multiplexer 440, and vice-versa.

Each tap multiplexer 440 is also coupled to an optical receiver 370. From the bi-directional splitter 360, respective optical receivers 370 can convert the upstream optical signals into the electrical domain. Each optical receiver 370 can comprise one or more photoreceptors or photodiodes that convert optical signals into electrical signals. Since

the optical transmitters **325** and optical receivers **370** can comprise off-the-shelf hardware to generate and receive respective optical signals, the transceiver node **120** lends itself to efficient upgrading and maintenance to provide significantly increased data rates.

Each optical transmitter **325** and each optical receiver **370** are connected to a  
5    respective bi-directional splitter **360**. Each bi-directional splitter **360** in turn is connected to a diplexer **420** which combines the unidirectional optical signals received from the splitter **415** with the downstream optical signals received from respective optical receivers **370**. In this way, broadcast video services as well as data services can be supplied with a single optical waveguide such as a distribution optical waveguide **150** as illustrated in  
10    Figure 2. In other words, optical signals can be coupled from each respective diplexer **420** to a combined signal input/output port **445** that is connected to a respective distribution optical waveguide **150**.

Unlike the conventional art, the transceiver node **120** does not employ a  
conventional router. The components of the transceiver node **120** can be disposed within  
15    a compact electronic packaging volume. For example, the transceiver node **120** can be designed to hang on a strand or fit in a pedestal similar to conventional cable TV equipment that is placed within the "last" mile or subscriber proximate portions of a network. It is noted that the term, "last mile," is a generic term often used to describe the last portion of an optical network that connects to subscribers.

Also because the optical tap routing device **435** is not a conventional router, it  
20    does not require active temperature controlling devices to maintain the operating environment at a specific temperature. In other words, the transceiver node **120** can operate in a temperature range between minus 40 degrees Celsius to 60 degrees Celsius in one exemplary embodiment.

While the transceiver node **120** does not comprise active temperature controlling  
25    devices that consume power to maintain temperature of the transceiver node **120** at a single temperature, the transceiver node **120** can comprise one or more passive temperature controlling devices **450** that do not consume power. The passive temperature controlling devices **450** can comprise one or more heat sinks or heat pipes that remove  
30    heat from the transceiver node **120**. Those skilled in the art will appreciate that the

present invention is not limited to these exemplary passive temperature controlling devices. Further, those skilled in the art will also appreciate the present invention is not limited to the exemplary operating temperature range disclosed. With appropriate passive temperature controlling devices 450, the operating temperature range of the transceiver node 120 can be reduced or expanded.

In addition to the transceiver node's 120 ability to withstand harsh outdoor environmental conditions, the transceiver node 120 can also provide high speed symmetrical data transmissions. In other words, the transceiver node 120 can propagate the same bit rates downstream and upstream to and from a network subscriber. This is yet another advantage over conventional networks, which typically cannot support symmetrical data transmissions as discussed in the background section above. Further, the transceiver node 120 can also serve a large number of subscribers while reducing the number of connections at both the data service hub 110 and the transceiver node 120 itself.

The transceiver node 120 also lends itself to efficient upgrading that can be performed entirely on the network side or data service hub 110 side. That is, upgrades to the hardware forming the transceiver node 120 can take place in locations between and within the data service hub 110 and the transceiver node 120. This means that the subscriber side of the network (from distribution optical waveguides 150 to the subscriber optical interfaces 140) can be left entirely in-tact during an upgrade to the transceiver node 120 or data service hub 110 or both.

Referring now to Figure 4, this Figure is a functional block diagram illustrating an optical tap 130 connected to a subscriber optical interface 140 by a single optical waveguide 150 according to one exemplary embodiment of the present invention. The optical tap 130 can comprise a combined signal input/output port 505 that is connected to another distribution optical waveguide that is connected to a transceiver node 120. As noted above, the optical tap 130 can comprise an optical splitter 510 that can be a 4-way or 8-way optical splitter. Other optical taps having fewer or more than 4-way or 8-way splits are not beyond the scope of the present invention. The optical tap can divide downstream optical signals to serve respective subscriber optical interfaces 140. In the

exemplary embodiment in which the optical tap 130 comprises a 4-way optical tap, such an optical tap can be of the pass-through type, meaning that a portion of the downstream optical signals is extracted or divided to serve a 4-way splitter contained therein, while the rest of the optical energy is passed further downstream to other distribution optical waveguides 150.

The optical tap 130 is an efficient coupler that can communicate optical signals between the transceiver node 120 and a respective subscriber optical interface 140. Optical taps 130 can be cascaded, or they can be connected in a star architecture from the transceiver node 120. As discussed above, the optical tap 130 can also route signals to other optical taps that are downstream relative to a respective optical tap 130.

The optical tap 130 can also connect to a limited or small number of optical waveguides so that high concentrations of optical waveguides are not present at any particular transceiver node 120. In other words, in one exemplary embodiment, the optical tap can connect to a limited number of optical waveguides 150 at a point remote from the transceiver node 120 so that high concentrations of optical waveguides 150 at a transceiver node can be avoided.

The subscriber optical interface 140 functions to convert downstream optical signals received from the optical tap 130 into the electrical domain that can be processed with appropriate communication devices. The subscriber optical interface 140 further functions to convert upstream electrical signals into upstream optical signals that can be propagated along a distribution optical waveguide 150 to the optical tap 130. The subscriber optical interface 140 can comprise an optical diplexer 515 that divides the downstream optical signals received from the distribution optical waveguide 150 between a bi-directional optical signal splitter 520 and an analog optical receiver 525. The optical diplexer 515 can receive upstream optical signals generated by a digital optical transmitter 530. The digital optical transmitter 530 converts electrical binary/digital signals to optical form so that the optical signals can be transmitted back to the data service hub 110. Conversely, the digital optical receiver 540 converts optical signals into electrical binary/digital signals so that the electrical signals can be handled by processor 550.

The present invention can propagate the optical signals at various wavelengths. However, the wavelength regions discussed are practical and are only illustrative of exemplary embodiments. Those skilled in the art will appreciate that other wavelengths that are either higher or lower than or between the 1310 and 1550 nm wavelength regions  
5 are not beyond the scope of the present invention.

The analog optical receiver **525** can convert the downstream broadcast optical video signals into modulated RF television signals that are propagated out of the modulated RF unidirectional signal output **535**. The modulated RF unidirectional signal output **535** can feed to RF receivers such as television sets (not shown) or radios (not  
10 shown). The analog optical receiver **525** can process analog modulated RF transmission as well as digitally modulated RF transmissions for digital TV applications.

The bi-directional optical signal splitter **520** can propagate combined optical signals in their respective directions. That is, downstream optical signals entering the bi-directional optical splitter **520** from the optical the optical diplexer **515**, are propagated to  
15 the digital optical receiver **540**. Upstream optical signals entering it from the digital optical transmitter **530** are sent to optical diplexer **515** and then to optical tap **130**. The bi-directional optical signal splitter **520** is connected to a digital optical receiver **540** that converts downstream data optical signals into the electrical domain. Meanwhile the bi-directional optical signal splitter **520** is also connected to a digital optical transmitter **530**  
20 that converts upstream electrical signals into the optical domain.

The digital optical receiver **540** can comprise one or more photoreceptors or photodiodes that convert optical signals into the electrical domain. The digital optical transmitter can comprise one or more lasers such as the Fabry-Perot (F-P) Lasers, distributed feedback lasers, and Vertical Cavity Surface Emitting Lasers (VCSELs).

25 The digital optical receiver **540** and digital optical transmitter **530** are connected to a processor **550** that selects data intended for the instant subscriber optical interface **140** based upon an embedded address. The data handled by the processor **550** can comprise one or more of telephony and data services such as an Internet service. The processor **550** is connected to a telephone input/output **555** that can comprise an analog interface. The  
30 processor **550** is also connected to a data interface **560** that can provide a link to computer

devices, set top boxes, ISDN phones, and other like devices. Alternatively, the data interface 560 can comprise an interface to a Voice over Internet Protocol (VoIP) telephone or Ethernet telephone. The data interface 560 can comprise one of Ethernet's (10BaseT, 100BaseT, Gigabit) interface, HPNA interface, a universal serial bus (USB) an  
5 IEEE1394 interface, an ADSL interface, and other like interfaces.

Referring now to Figure 5, this figure illustrates a functional block diagram of an exemplary optical tap routing device 435 and a tap multiplexer 440. This figure further illustrates the exemplary hardware that can be found in each tap multiplexer 440. However, those skilled in the art will recognize the present invention is not limited to the  
10 hardware illustrated nor is the present invention limited to a hardware embodiment. That is, software or other hardware or a combination thereof can be substituted for the elements described in Figure 5 without departing from the scope and spirit of the present invention.

For downstream communications signals, the optical tap routing device 435 can route or divide or apportion data service hub signals according to the individual tap  
15 multiplexers 440 that communicate optical signals with one or more optical taps 130 and ultimately one or more subscriber optical interfaces 140 (not shown in Figure 5). In the downstream direction, it is noted that tap multiplexer 440 receives electrical signals from the optical tap routing device 435. That is, the tap multiplexer 440 operates in the electrical domain to modulate laser transmitters in order to generate optical signals that  
20 are assigned to groups of subscribers coupled to one or more optical taps. The optical tap routing device 435, as noted above, can comprise a computer or hardwired apparatus that executes a program defining a protocol for communications with groups of subscribers assigned to individual ports. The optical tap routing device can assign multiple subscribers to a single port. More specifically the optical tap routing device can service  
25 groups of subscribers with corresponding respective, single ports. Attached to each port of the optical tap routing device 435 are tap multiplexer 440.

Tap multiplexer 440 can propagate optical signals to and from various groupings of subscribers. In one exemplary embodiment, a tap multiplexer 440 can comprise classifiers 562, policers 564, and a plurality of priority output buffers 566, 568, 570 and  
30 572. A tap multiplexer 440 receives downstream data packets from the optical tap routing

device 435. The classifiers 562 identify these outbound packets (outbound relative to the data service hub 110 of the optical network) and assign each packet to an appropriate class. In other words, each classifier 562 can select a packet based on the content of packet headers according to predefined rules.

- 5       Classes can be defined by the values of arbitrary bits in the packet header, and each classifier 562 can examine up to 40 bytes (or 320 bits) of each packet. Each classifier 562 can consider multiple fields of an individual packet, including the full Ethernet header, the full IP header, and the source and destination TCP or UDP ports. The Ethernet header can comprise a destination media access control (MAC) address as  
10       well as a source MAC address. Other headers available for classification include, but are not limited to, those fields listed in Table 1 below.

Table 1. - Header Fields Available for Classification

Ethernet Header (14 bytes)	destination MAC address					
	source MAC address					
	Ethernet type					
IP Header (20 bytes)	vers	hlen	diffserv	ECN	payload length	
	fragment identifier			0 DM FF	fragment offset	
	hop limit		next header		header checksum	
	source IP address					
	destination IP address					
Partial UDP/TCP Header (6 bytes)	source UDP/TCP port			destination UDP/TCP port		
	UDP message length					

15

In one exemplary embodiment, the tap multiplexer **440** can comprise a plurality of separate classifiers **562** for each logical channel that supports a preassigned grouping of subscribers. That is, in one exemplary embodiment, each logical channel can support sixteen different subscribers. However, the present invention is not limited to this particular number of subscribers per logical channel. A fewer or an increased amount of subscribers can be assigned to each logical channel without departing from the scope and spirit of the present invention. Each classifier **562** can be configured with the following values: A 40-byte bit mask; a 40-byte check value; and a policer assignment.

Each policer **564** can be coupled to a corresponding classifier **562**. However, in an alternative embodiment (not illustrated), multiple classifiers **562** may be coupled to a single policer **564**. Each policer **564** may operate as a two-stage token bucket where the first stage bucket can enforce a configured peak rate for the down stream communication traffic. Peak rate can comprise the maximum rate that a subscriber (via a subscriber optical interface **140**) is allowed to transmit downstream packets. Specifically, it may comprise the maximum rate at which the network will accept traffic bursts from the subscriber (via a subscriber optical interface **140**), expressed in bits per second. At this first stage, non-conforming packets that do not match the peak rate set in a policer **564** can be discarded.

The second stage of each traffic policer **564** operating as a token bucket can identify packets that conform to a sustained rate. Sustained rate can comprise the minimum throughput that the network will provide to the user, expressed in bits per second (Bps). At the second stage of each policer **564**, a burst size can also be evaluated. Burst size usually comprises the amount of traffic that the network will accept without pause at the user's peak rate, expressed in bits.

The classifiers **562** and policers **564** can comprise hardware such as applications specific integrated circuits (ASICs) or field programmable gate arrays (FPGAs). While the classifiers **562** and policers **564** may comprise ASICs or FPGAs, the present invention is not limited to these hardware devices. Other similar processing devices are not beyond the scope of the present invention. Further, as noted above, the present invention is not



limited to the hardware illustrated and can also be embodied in software or a combination thereof, without the departing from the scope and spirit of the present invention.

In one exemplary embodiment, the classifiers **562** can distinguish different traffic classes based on the differentiated services code point (DSCP) in each packet's header.

5 DSCP values are defined in RFC 2474, published by the internet engineering task force (IETF) available at the web site [www.ietf.org](http://www.ietf.org). The six bits of the DSCP value is the successor to the so called "precedence" bits defined in RFC 791. The precedence definition is modified and expanded in RFC 2474. The relevant bits of the DSCP values are sometimes referred to as ToS (Type of Service) bits in IPv4 (the version of Internet

10 Protocol most commonly used as of the filing of this document) and are called the traffic class octet in IPv6 (a newer version of the Internet Protocol not in widespread use on the public internet as of the filing date of this document).

Once the classifiers **562** have identified traffic with the desired DSCP values (or other parameters as described later in this description), they can pass the traffic to the

15 appropriate policer **564**. The policers **564** enforce a maximum transmission rate (also referred to as the peak rate), a minimum transmission rate (also referred to as the sustained rate), and maximum burst size for the downstream communication traffic. If the downstream traffic exceeds the maximum transmission rate, excess packets above that maximum transmission rate are discarded. If the downstream traffic exceeds the minimum

20 transmission rate, excess traffic above that minimum transmission rate is marked as "out of profile."

The classifiers **562** can use DSCP values (or other parameters as mentioned later in this description) to determine the policer assignment and ultimately which priority buffer will handle a particular packet. As noted above, each policer **564** is associated with

25 a particular output buffer that has a preset priority relative to other output buffers. The higher the priority buffer, the sooner or earlier the packet will be transmitted when more than one packet is ready for transmission to the subscribers because packets placed in higher priority output buffers are transmitted before packets in lower priority output buffers. By transmitting packets with high priority first, these packets have first access to

30 the guaranteed bandwidth, meaning that they will be handled immediately, assuming

adequate bandwidth is available.

Each priority output buffer **566**, **568**, **570**, and **572** can comprise a first-in/first-out register (FIFO). However, the buffers of the present invention are not limited to FIFO registers. Other memory devices that function similar to FIFOs are not beyond the scope of the present invention. Further, the present invention is not limited to the number of buffers illustrated. More or fewer buffers could be used without departing from the scope of the present invention.

Referring now to Figure 5, as each packet enters from the optical tap routing device **435**, it is identified by one of the classifiers **562**, based on a number of parameters that can be set by the operator. These parameters can comprise DSCP code values among other things as will be discussed below. An appropriate classifier **562**, if any, selects the packet. Each classifier has a particular policer assignment that is given to a packet. Through the policer mapping function **631**, the packet is transferred or mapped to the appropriate policer **564** based upon the policer assignment given by the classifier **562**. More than one classifier **562** can assign packets to the same policer **564**, but one classifier **562** usually may not assign packets to more than one policer. During a first stage (i.e., a first token bucket algorithm) of a policer **564**, it determines if the packet is within an allowable peak data rate, as determined by its classification. If not, the packet is dropped. If the packet is within the allowed peak data rate then the policer's second stage (i.e., a second token bucket algorithm) determines if the packet is within a guaranteed or sustained rate and if the packet is within a burst size. All packets, whether or not they are within the guaranteed rate or burst size, are passed to one of the output buffers **566**, **568**, **570** or **572** via an output buffer mapping function **665**. Each policer **564** passes packets to one output buffer. Any policer **564** may pass packets to any output buffer, but can usually pass packets only to one output buffer. The output buffer to which a particular policer passes packets is usually determined by the network service provider when he sets up his data traffic policies.

As noted above, one distinguishing feature of the policers **564** of the present invention is their relative physical location within the optical network as well as the type of data traffic that each policer **564** handles. As is understood to those skilled in the art,

5      policers typically function at a network border (an ingress point) that ensures that a host does not violate its promised traffic characteristic. Policers of the conventional art typically limit the amount of traffic flowing into a network to achieve a specific policy goal. Policers of the conventional art typically monitor and control traffic as the traffic enters the network. However, according to the present invention, the policers **564** are employed within tap multiplexers **440** that are in close proximity to the subscribers.

10      Policers **564** of the present invention function at a network border, but at egress points rather than ingress points, compared to that of the conventional art. In this way, the policers **564** can control the volume or content (or both) of downstream communications that exit an optical network that are received by subscribers of the optical network. The control of volume or content (or both) is a result of the policers **564** evaluating the peak rate, sustained rate, and burst size of a packet. This control can also be attributed to a policer **564** assigning a packet with a particular weighted random early discard value. Those skilled in the art appreciate that Internet traffic can be slowed down  
15      if packets are dropped, so that if packets to a particular destination are being dropped, then eventually the rate at which packets leave the optical network of the present invention towards a destination (such as a subscriber) may be reduced.

20      As noted above, each policer **564** can be configured with the following exemplary values: a peak rate, a profile rate, a burst size, Weighted Random Early Discard (WRED) parameters for in-profile traffic, WRED parameters for out-of-profile traffic, and next stage output buffer assignment. While the burst size can comprise the amount of data the subscriber can receive at its peak rate without pause or delay, expressed in bits, the burst size can also comprise a special value to indicate that a subscriber has no limit on his or her burst size. The WRED parameters will be discussed in further detail below with  
25      respect to Figures 6 through 10.

30      Each output buffer **566**, **568**, **570**, and **572** takes in packets after a respective buffer executes the weighted random early discard algorithm as each packet is presented to a particular buffer. Each output buffer can then send the packet downstream if that particular buffer is requested to release its stored packets. The first priority output buffer **566** can evaluate all packets which have been determined to have the highest priority, and

hence should be transmitted first towards the subscribers during downstream processing. Successive output buffers have lower priority down to the lowest priority fourth output buffer 572.

As mentioned above, each priority output buffer separately implements a  
5 Weighted Random Early Discard (WRED) algorithm to determine if packets are admitted to the buffer or dropped. Each priority output buffer operates differently for traffic that conforms to the values assigned to a policer and for downstream traffic that does not conform to the values assigned to a particular policer.

Specifically, downstream traffic that is considered within preset parameters  
10 assigned to a policer by a network service provider (such as peak rate, sustained rate, and burst size) is subject to a Weighted Random Early Discard algorithm according to three parameters: A minimum threshold, a maximum threshold, and a maximum drop probability that is specific to in-profile traffic. The minimum threshold, maximum threshold, and maximum drop probability are assigned to each policer 564 by a network  
15 service provider.

For downstream traffic falling outside of a policer's preset parameters, this traffic is also subject to a Weighted Random Early Discard (WRED) algorithm according to three parameters: a minimum threshold, a maximum threshold, and a maximum drop probability that is specific to out-of-profile traffic and also assigned by each policer 564.  
20 As noted above, the minimum threshold, maximum threshold, and maximum drop probability are assigned to each policer 564 by a network service provider.

By using different values for the maximum drop probability for traffic falling within and outside a policer's preset values, this allows different traffic classes to be weighted differently. In effect, the service provider may assign traffic priority according  
25 to a WRED algorithm.

Once packets are stored in a particular priority output buffer, the packets are removed from each respective priority output buffer according to a predetermined policy or queuing discipline. Typically, packets are removed from any particular output buffer only when all higher priority output buffers are empty. For example, if packets are  
30 present in each of the priority output buffers 566, 568, 570 and 572, packets in the second

priority output buffer 568 would not start being removed until all of the packets in the first priority output buffer 566 are removed. Similarly, packets stored in the third priority output buffer 570 would not be removed for downstream communications until all of the packets in the second priority output buffer 568 are removed. Such a queuing discipline or output buffer policy provides lower delay for high priority downstream traffic.

Referring now to Figure 6, this figure illustrates an exemplary method for handling downstream communications originating from a data service hub 110 of an optical network that are transmitted to subscribers of the optical network. Basically, Figure 6 provides an overview of the processing performed by the optical tap routing device 435 and tap multiplexer 440 housed within the transceiver node 120.

The description of the flow charts that follows is represented largely in terms of processes and symbolic representations of operations by conventional computer components, including a processing unit (a processor), memory storage devices, connected display devices, and input devices. Furthermore, these processes and operations may utilize conventional computer components in a heterogeneous distributed computing environment, including remote file servers, computer servers, and memory storage devices. Each of these conventional distributed computing components can be accessible by the processor via a communication network.

The processes and operations performed below may include the manipulation of signals by a processor and the maintenance of these signals within data structures resident in one or more memory storage devices. For the purposes of this discussion, a process is generally conceived to be a sequence of computer-executed steps leading to a desired result. These steps usually require physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical, magnetic, or optical signals capable of being stored, transferred, combined, compared, or otherwise manipulated. It is convention for those skilled in the art to refer to representations of these signals as bits, bytes, words, information, elements, symbols, characters, numbers, points, data, entries, objects, images, files, or the like. It should be kept in mind, however, that these and similar terms are associated with appropriate physical quantities

for computer operations, and that these terms are merely conventional labels applied to physical quantities that exist within and during operation of the computer.

It should also be understood that manipulations within the computer are often referred to in terms such as creating, adding, calculating, comparing, moving, receiving, determining, identifying, populating, loading, executing, etc. that are often associated with manual operations performed by a human operator. The operations described herein can be machine operations performed in conjunction with various input provided by a human operator or user that interacts with the computer.

In addition, it should be understood that the programs, processes, methods, etc. described herein are not related or limited to any particular computer or apparatus. Rather, various types of general purpose machines may be used with the following process in accordance with the teachings described herein.

The logic flow described in Figure 6 can be the core logic or top level processing and can be executed repeatedly. The logic flow diagram illustrated in Figure 6 illustrates a process that can occur after initialization of the software or hardware components or both illustrated in Figures 1-5.

For example, in an object-oriented programming environment, software components or software objects or hardware that could be used to perform the steps illustrated in Figure 6 can be initialized or created prior to the process described in Figures 4 and 5. Therefore, one of ordinary skill in the art recognizes that several steps pertaining to initialization of software objects or hardware described in Figures 1 through 5 may not be illustrated.

The present invention may comprise a computer program or hardware or a combination thereof which embodies the functions described herein and illustrated in the appended flow charts. However, it should be apparent that there could be many different ways of implementing the invention in computer programming or hardware design, and the invention should not be construed as limited to any one set of computer program instructions. Further, a skilled programmer would be able to write such a computer program or identify the appropriate hardware circuits to implement the disclosed invention without difficulty based on the flow charts and associated description in the

application text, for example. Therefore, disclosure of a particular set of program code instructions or detailed hardware devices is not considered necessary for an adequate understanding of how to make and use the invention. The inventive functionality of the claimed computer implemented process will be explained in more detail in the following  
5 description in conjunction with the remaining Figures illustrating the process flow.

Certain steps in the processes or process flow described below must naturally precede others for the present invention to function as described. However, the present invention is not limited to the order of the steps described if such order or sequence does not alter the functionality of the present invention. That is, it is recognized that some  
10 steps may be performed before or after other steps without departing from the scope and spirit of the present invention.

Step 610 is the first step in the exemplary method 600 processing downstream communications. In step 610, a packet is received from the optical tap routing device 435 by tap multiplexer 440.

15 In decision step 615, it can be determined whether a packet matches more than one classifier 562 of a particular tap multiplexer 440. If the inquiry to decision step 615 is positive then the “yes” branch is followed to step 620 in which the packet is assigned to one of the matching classifiers 562 according to an order that can be established by the service provider. If the inquiry to decision step 615 is negative, then the “no” branch is  
20 followed to decision step 625.

In decision step 625, it is determined whether a packet matches any of the classifiers 562 of a particular tap multiplexer 440. If the inquiry to decision step 625 is negative, then the “no” branch is followed to step 630 in which the packet is dropped. If the inquiry to decision step 625 is positive, then the “yes” branch is followed to step 631.

25 In step 631, the packet is mapped to the appropriate policer 564 that is associated with the classifier 562 that previously processed the packet. As noted above, each classifier 562 is assigned to a single policer 564. Each policer 564 is typically associated with a single classifier 562 and a single priority output buffer.

In decision step 635, each respective policer 564 can determine whether a packet exceeds a peak rate for the destined subscriber. As noted above, peak rate can comprise the maximum rate that a subscriber is allowed to receive downstream packets. Specifically, it may comprise the maximum rate at which the network will accept traffic bursts bound to the user, expressed in bits per second. Decision step 635 is highlighted with a dashed routine symbol to indicate that it comprises a first stage token bucket algorithm for evaluating the peak rate for a subscriber. Those skilled in the art are familiar with token bucket algorithms. One reference which describes such bucket algorithms is the following publication: "Policing and Shaping Overview," published by Cisco Systems, Inc., pages QC 87- QC 98. Another exemplary publication describing token bucket algorithms is the following white paper: "Cisco IOS(TM) Software Quality of Service Solutions," published by Cisco Systems, Inc., copyright 1998. The contents of both these reference are incorporated fully herein by reference.

If the inquiry to decision step 635 is positive, then the "yes" branch is followed to step 637 in which the packet is dropped. If the inquiry to decision step 635 is negative, then the "no" branch is followed to decision step 640.

In decision step 640, a policer 564 can determine if a packet matches a sustained rate and burst size. Decision step 640 is also highlighted with a dashed routine symbol to indicate that it comprises a second stage token bucket algorithm for evaluating the peak rate for a subscriber. As noted above, those skilled in the art are familiar with token bucket algorithms and therefore, a detailed discussion of these algorithms will not be provided. The reader is referred to the aforementioned token bucket algorithm publications which are fully incorporated herein by reference. If the inquiry to decision step 640 is negative, then the "no" branch is followed to step 645 in which the packet is identified as non-conforming with burst size or sustained rate assigned to the policer 564 by a network administrator. Next, in step 650 the policer 564 can assign a "non-conforming" maximum drop probability, a maximum threshold, and minimum threshold to the packet that is specific to traffic that is determined as "out-of-profile" meaning that the packet is outside (greater than) the policer's burst size or sustained rate.



If the inquiry to decision step 640 is positive, then the "yes" branch is followed to step 655 in which the policer 564 can identify the packet as conforming with a traffic profile for a particular classifier 562. Next, in step 660, a policer 564 can assign a conforming maximum drop probability, a maximum threshold, and a minimum threshold  
5 to the packet that is specific to traffic that is determined as "in-profile" meaning that the packet is within the policer's burst size and sustained rate.

In step 665, the packet is mapped to the appropriate output buffer. Typically, each policer 564 is associated with a particular output buffer 566, 568, 570, and 572. In decision step 670, each priority output buffer can determine whether a packet is identified  
10 as either in-profile traffic or out-of-profile traffic. If the inquiry to decision step 670 is positive, meaning that a particular packet matches the burst size or sustained rate assigned to the policer then the "yes" branch is followed to routine 675 in which a particular output buffer determines whether to admit the conforming packet to the assigned output buffer. Further details of routine 675 will be discussed below with respect to Figure 7.

15 If the inquiry to decision step 670 is negative, meaning that a packet does not conform with the sustained rate or burst size assigned to a policer 564, then the "no" branch is followed to routine 680 in which the particular output buffer determines whether to admit the nonconforming packet to the assigned output buffer. Further details of routine 680 will be discussed below with respect to Figure 8.

20 In step 685, the packets admitted to the buffers are removed in a predetermined order as discussed above. Typically, this predetermined order comprises removing packets from higher priority buffers first and then removing packets from lower priority buffers last. In step 690, the packets are forwarded to the subscribers.

Referring now to Figure 7, this figure illustrates an exemplary subprocess 675 for  
25 determining whether to admit in-profile packets into a particular priority output buffer. This figure provides an overview of the processing performed by each of the priority output buffers.

Certain steps in the process described below must naturally proceed others for the present invention to function as described. However, the present invention is not limited  
30 to the order of the steps described in such order of sequence of steps does not alter the

functionality of the present invention. That is, it is recognized that some steps may be performed before or after other steps without departing from the scope and spirit of the present invention.

Step 705 is the first step in the exemplary method 675 for admitting in-profile  
5 packets to a particular priority output buffer. In step 705, it is determined whether the particular output buffer of interest is full. If the inquiry to decision step 705 is positive, then the “yes” branch is followed to step 710 in which the packet or series of packets are dropped. Then in step 720, the process returns to step 600 of Figure 6.

If the inquiry to decision step 705 is negative, then the “no” branch is followed to  
10 step 725 in which the receiving output buffer’s average fill or current volume is determined. In step 725, the output buffer’s average fill or average current volume is computed by only counting conforming packets. In other words, the output buffer’s average current volume is calculated based only upon those packets conforming with a particular communication traffic profile.

15 In decision step 730, it is determined whether the calculated output buffer average fill or volume is below a “conforming” minimum threshold. If the inquiry to decision step 730 is positive, then the “yes” branch is followed to step 735 in which the packet is stored in the output buffer. Next, in step 740, the process returns to step 685 of Figure 6.

If the inquiry to decision step 730, is then negative, then the “no” branch is  
20 followed to the decision step 745 in which it is determined whether the calculated output buffer average fill or volume is above a “conforming” maximum threshold. If the inquiry to decision step 745 is positive, then the “yes ” branch is followed step 750 in which the packet is dropped. The process then returns to step 615 of Figure 6.

If the inquiry to decision step 745 is negative, then the “no” branch is followed to  
25 step 760 in which the packet can be dropped according to a Weighted Random Early Discard (WRED) algorithm. The WRED algorithm typically uses an exponentially weighted moving average estimator to compute the average output buffer (queue) fill or volume which in turn typically smoothes out any bursty packet flow. The probability of packet drop typically increases as the average queue or buffer fill or volume increases. A  
30 packet is typically discarded with a probability that varies linearly from zero (when the

average buffer volume is at the minimum threshold) to the configured maximum drop probability (when the average buffer volume is at the maximum threshold). Figure 10 illustrates the WRED algorithm in a graphical fashion for in-profile or conforming downstream traffic.

- 5       The WRED algorithm uses an exponentially weighted moving average to calculate the average buffer size as discussed above. The measurement of the average buffer size is updated each time a packet is presented for admission to a particular priority output buffer or queue. The algorithm updates the average buffer size by using the previous value and an instantaneous value of the average buffer size, according to the equation listed below:

10       
$$Q_{avg} = ( 255 / 256 \cdot Q_{avg} ) + ( 1 / 256 \cdot Q_{inst} )$$

where  $Q_{avg}$  is the average buffer size; and  $Q_{inst}$  is the instantaneous average buffer size.

- As Figure 10 illustrates, when the average buffer size or queue depth is above a minimum threshold ( $Th_{min}$ ), the WRED algorithm starts dropping packets. The rate of packet drop typically increases linearly as the average buffer or queue fill/volume increases until the average queue size reaches a maximum threshold ( $Th_{max}$ ). In Figure 10,  $P_{max}$  denotes the maximum drop probability assigned to the current packet by a policer 564.

- Referring now to Figure 8, this figure illustrates an exemplary subprocess 680 for admitting out-of-profile packets to a particular buffer. This figure provides an overview of processing performed by priority output buffers for out-of-profile packets.

- Certain steps in the process described below must naturally proceed others for the present invention to function as described. However, the present invention is not limited to the order of steps described in such order or sequence does not alter the functionality of the present invention. That is, it is recognized that some steps may be performed before or after other steps without departing from the scope and spirit of the present invention.

Step 805 is the first step in exemplary subprocess 680 of admitting out-of-profile packets to a priority output buffer. In step 805, it is determined whether the particular output buffer of interest is full. If the inquiry to decision step 805 is positive, then the

“yes” branch is followed to step 810 in which the packet is dropped. Next, in step 815, the process returns to step 605 of Figure 6.

If the inquiry to decision step 805 is negative, then the “no” branch is followed to step 820 in which the output buffer's average fill or current volume is calculated. In step 5 820, the output buffer's average volume is calculated by counting both conforming and non-conforming packets.

In decision step 825, it is determined whether the output buffer average fill or volume is below a “non-conforming” minimum threshold. If the inquiry to decision step 825 is positive, then the “yes” branch is followed to step 830 in which the packets are 10 stored in the output buffer. Next, in step 835, the process returns to step 605 of Figure 6.

If the inquiry to decision step 825 is negative, then the “no” branch is followed to decision step 840 in which it is determined whether the output buffer average fill or volume is above a “non-conforming” maximum threshold. If the inquiry to decision step 840 is positive, then the “yes” branch is followed to step 845 in which the packet or series 15 of packets are dropped. In step 850, the process returns to step 605 of Figure 6.

If the inquiry to decision step 840 is negative, then the “no” branch is followed to step 855 in which the one or more packets are dropped according to a Weighted Random Early Discard algorithm (WRED), as discussed above. However, the WRED algorithm for step 855 uses different parameters than does the WRED algorithm of step 760 of 20 Figure 7. The difference lies in the maximum probability drop value ( $P_{\max}$ ) and the minimum and maximum threshold values  $Th_{\max}$  and  $Th_{\min}$ . See Figure 10 for definitions of terms. As noted above with respect to step 820 of subprocess 680, an output buffer's average fill or volume is computed counting both conforming and non-conforming packets.

25 On the other hand, in step 725 of Figure 7, an output buffer's average fill or volume is computed counting only conforming packets which match the communication traffic profile rate for a particular subscriber. Another difference exists in the threshold values assigned to in-profile traffic and out-of-profile traffic. The threshold values for in-profile traffic are different from those of out-of-profile traffic.

By using multiple values for the maximum drop probability as well as adjusting the threshold values for in-profile traffic and out-of-profile traffic, specific traffic classes can be weighted differently. In effect, such a feature lets a service provider assign traffic priority over other types of traffic. As long as the output buffer size is between the configuration thresholds, the probability of a packet being dropped is directly proportional to the maximum drop probability that the service provider assigns to it. As Figure 9 illustrates (compared to Figure 10), the threshold values for  $Th_{min}$ ,  $Th_{max}$ , for out of profile traffic are generally lower than for in-profile traffic, and the maximum drop probability is higher for this out-of-profile traffic.

## IMPLEMENTING DOWNSTREAM QoS POLICY

The present invention allows service providers to define powerful and flexible quality of service management rules. The following describes how to use those rules in practice. Several aspects of QoS policy, including, but not limited to, prioritization, mapping of backbone priorities, and subscriber bandwidth limitations can be implemented with the present invention.

### Voice Traffic

In many environments, some traffic may be given higher priority than others. Voice over IP and TDM over IP packets, for example, can benefit if given priority over normal data traffic. Both of these traffic types are destined for the subscriber optical interfaces (SOIs), rather than for subscriber equipment attached to the Subscriber Optical interfaces.

To ensure that this traffic receives an appropriate priority, it can be assigned to one or more classifiers. Since all such packets typically have the SOI itself as the IP and MAC destination, one convenient classification relies on the IEEE Organizationally Unique Identifier (OUI) in the destination MAC address. In one exemplary embodiment, these three bytes can have the value  $00060D_{16}$ .

The subscript 16 of the previous value indicates that the number is expressed in that base. Similarly other numbers are expressed in base 2 and in base 10, and are

similarly identified to reduce any possible confusion. These bases are well understood by those skilled in the art. All mask and values shown below are understood to be expressed in base 16. The classifier mask and value, therefore, can be set to the following values:

5       Mask 1: FFFFFFFF000000000000000000000000  
           00000000000000000000000000000000  
           000000000000  
  
           Value 1: 00060D000000000000000000000000  
 10       00000000000000000000000000000000  
           000000000000

Those skilled in the art understand the mask and value to correspond to the sections of Table 1 of this description. Each character represents four bits of the corresponding value in table 1, expressed in base 16. Thus, each character in the mask and the value represent four bits of the four bytes (32 bits) occupying space from left to right in each row of Table 1. The first line of the mask represents the Ethernet header (14 bytes, so 28 characters in the mask and value). The next line represents the 20 bytes of the IP header of Table 1, and the last row represents the partial UDP/TCP header (6 bytes).

When the base 16 characters of the mask are converted to binary, a binary 1 represents a bit position that will be tested by the value, and a binary 0 represents a bit position that will not be checked. When the "value" characters are converted to binary, all "value" bit positions where there is a binary 1 in the mask, usually must be the same as the corresponding bit in the packet header, for the a classifier to accept the packet. If one or more of the bits are not the same, then the packet does not meet that classification, and drops to the next classifier. If it matches none of the classifiers, it is dropped. This is understood by those skilled in the art.

A single policer 564 can manage the bandwidth for the traffic represented by mask 1 and value 1. This is true even if a plurality of subscribers are receiving this type of traffic.

A typical residential deployment will support voice calls but not TDM over IP. Each voice call usually requires about 156.8 kbit/s of bandwidth. (This bandwidth

assumes G.711 codec and 5 ms sampling interval. Bandwidth includes the RTP, UDP, IP, and MAC headers and trailers, but not the preamble or inter-frame gap.)

For this example, assume the policer **564** needs to consider up to two simultaneous calls for each of 16 subscribers, plus allowance for other traffic to the SOI  
5 **140** (e.g. network management). The total bandwidth requirement is about 6 Mbit/s.

Since voice traffic is typically a constant bit rate, little burst capability is needed. Assume, as a worst case, that two samples for each call arrive consecutively. At 784 bits per packet, that would likely represent a burst of just over 25 kbit. Doubling this value to allow for network management and other overhead yields a burst limit of 50 kbit.

10 The policer **564** for this traffic, therefore, may be configured as follows:

Peak Rate 1: 9 Mbit/s

Profile Rate 1: 6 Mbit/s

Burst Limit 1: 50 kbit

Since voice traffic is particularly delay sensitive, it may be assigned to the highest  
15 output buffer or first priority output buffer **566**.

The peak rate 1 above is related to the first stage of the token bucket in the policer **564**. That first stage token bucket in step 635 in Figure 6 would be set to 9 Mbit/s by having tokens added at that rate. The profile rate 1 represents the second stage token  
20 bucket (step 640), which token bucket is filled at the rate corresponding to 6 Mbit/s. The burst limit determines how much data can pass at one time, and is the number of tokens in the second stage token bucket. In the example, the second stage token bucket can hold a maximum number of tokens representing 50 kbits of data.

## 25 Mapping Backbone Priorities

If a service provider uses, for example, diffserv code points to mark high priority traffic on its backbone, a similar approach can be used to prioritize traffic across the Optical Network. The expedited forwarding (EF) per hop behavior (PHB), for example, uses the diffserv code point value of 101110<sub>2</sub>. A classifier can be easily defined to identify  
30 this traffic.

Mask 2: 000000000000000000000000FFFF  
00FC000000000000000000000000000000  
000000000000

5

Value 2: 00000000000000000000000000800  
00B8000000000000000000000000000000  
000000000000

10 As an example, assume that expedited forwarding traffic is limited to 1000 Mbit/s,  
with normal rates of 100 Mbit/s and bursts up to 1 second in duration.

Peak Rate 2: 1000 Mbit/s

Profile Rate 2: 100 Mbit/s

15 Burst Limit 2: 100 Mbit

Since expedited forwarding presumes high priority, this traffic may be assigned to  
the highest priority output buffer or first priority output buffer 566. (This output buffer  
can be the same as used for voice and TDM traffic as discussed above.)

20

### Blocking Applications

Service providers may wish to completely block specific applications from their  
network. One way to do that is to assign those applications zero bandwidth. Consider, as  
an example, a provider that wishes to ban Napster traffic (Digital Music file sharing or  
25 other bulk file transfers) on its network. Napster servers typically use ports 7777<sub>10</sub>,  
8875<sub>10</sub>, and 8888<sub>10</sub>, so identifying all traffic from Napster servers can require three  
classifiers. Note that these classifiers, in addition to looking at TCP port numbers can  
also ensure that the datagrams (the data contained in the packets) are not fragments, other  
than the first of two packets across which one longer datagram was fragmented. This is  
30 understood by those skilled in the art.



Mask 3: 00000000000000000000FFFF  
0F0000000001FFF0FF0000000000000000  
FFFF00000000

5 Value 3: 000000000000000000000800  
0500000000000000006000000000000000  
1E6100000000

10 Mask 4: 00000000000000000000FFFF  
0F0000000001FFF0FF0000000000000000  
FFFF00000000

15 Value 4: 000000000000000000000800  
0500000000000000006000000000000000  
22AB00000000

20 Mask 5: 00000000000000000000FFFF  
0F0000000001FFF0FF0000000000000000  
FFFF00000000

Value 5: 000000000000000000000800  
0500000000000000006000000000000000  
22B800000000

25 All three of these classes can be assigned to a single policer. It is noted that this is an example of three classifiers **562** supplying packets to a single policer **564**. The bandwidth assignment is straightforward.

Peak Rate 3: 0 Mbit/s

Profile Rate 3: 0 Mbit/s

30 Burst Limit 3: 0 Mbit

The priority queue assignment for this traffic is irrelevant. For convenience, it may be assigned the lowest priority queue or fourth priority output buffer **572**.

#### Rate Limiting Traffic Types

35 The present invention can also limit the bandwidth of particular traffic types. For example, a service provider may wish to limit multicast streaming to 200 Mbit/s across all subscribers on a logical channel. Multicast traffic has an IP destination address whose first four bits are 1110<sub>2</sub>, and the Real Time Streaming Protocol (used as the basis for

Apple QuickTime and Real Networks RealVideo) typically uses destination port 554. To identify multicast RTSP packets, the following exemplary classifier configuration can be used:

5           Mask 6: 000000000000000000000000FFFF  
            0F00000000001FFF00FF000000000000F0000000  
            0000FFFF0000

            Value 6: 0000000000000000000000000800  
10           0500000000000000000011000000000000E0000000  
            0000022A0000

The rate governor for this traffic may be configured for 200 Mbit/s with a burst limit of 1.5 seconds.

15           Peak Rate 4: 250 Mbit/s (note that this exemplary peak rate is arbitrary, since speeds over 200 Mbits/s are not to be allowed.)

            Profile Rate 4: 200 Mbit/s

            Burst Limit 4: 300 Mbit

            Streaming applications are somewhat delay sensitive, so it may be beneficial to  
20           assign this traffic the second highest priority or second priority output buffer 568.

#### Protecting Against Denial of Service Attacks

A common type of denial of service attack relies on flooding the victim with ICMP Internet Control Message Protocol (ICMP) - used for internal housekeeping on the  
25           Internet) requests. Since legitimate uses of ICMP diagnostics require only a small amount of bandwidth, limiting the rate of ICMP traffic can protect against ICMP-based denial of service attacks. ICMP messages usually have a protocol value of 1 in the IP header.

30           Mask 7: 000000000000000000000000FFFF  
            000000000000000000FF0000000000000000000  
            000000000000

5      Peak Rate 5: 256 Kbit/s  
      Profile Rate 5: 256 Kbit/s  
      Burst Limit 5: 0 bit

## Prioritizing Premium Services

```
Value 8: 00000000000000000000000000800  
05000000000000000000600000000000000000  
06BB00000000
```

```

35      Value 9: 00000000000000000000000000800
      000000000000000000002F000000000000000000
      000000000000

```

Mask 10: 000000000000000000000000FFFF  
0F0000000001FFF00FF000000000000000000  
FFFF00000000

5

Value 10: 0000000000000000000000000800  
05000000000000000001100000000000000000  
01F400000000

10

Mask 11: 000000000000000000000000FFFF  
0F0000000001FFF00FF000000000000000000  
FFFF00000000

15

Value 11: 0000000000000000000000000800  
05000000000000000001100000000000000000  
06A500000000

The peak and profile rates for each subscriber may be assigned according to the service level agreement.

20

#### Subscriber Bandwidth Assignments

A key feature of the present invention is detailed management of bandwidth assigned to each subscriber. The flexibility offered by the present invention system in this area is nearly unlimited; the following merely shows a representative example.

25

For this exemplary embodiment, the service provider can define three levels of service for Internet access—premium, standard, and entry. The entry-level service can be roughly comparable to existing cable modem and digital subscriber line (DSL) services. It can offer 1 Mbit/s of bandwidth and best-effort delivery. The standard service can provide Ethernet-equivalent performance: 10 Mbit/s of bandwidth and best-effort delivery. The premium service can double the bandwidth—to 20 Mbit/s—and it can offer priority delivery. Premium traffic can be prioritized ahead of standard and entry-level traffic.

30

With such service definitions the QoS configuration can be relatively straightforward. Traffic classifiers can match the destination IP subnetwork of each

35

subscriber. For example, suppose that 16 subscribers are each given 28-bit subnetworks from the 10.0.0.0 range. (Subscriber 1 is 10.0.0.0/28, subscriber 2 is 10.0.0.16/28, and so on, all the way to 10.0.0.240/28. The /28 indicates that only the first 28 bits of the address are represented.) A total of 16 classifiers is needed to distinguish all 16 subscribers:

5		Mask 12: 000000000000000000000000FFFF 00000000000000000000000000000000FFFFF0 000000000000
10		Value 12: 0000000000000000000000000800 00000000000000000000000000000000A000000 000000000000
15		Mask 13: 000000000000000000000000FFFF 00000000000000000000000000000000FFFFF0 000000000000
20		Value 13: 0000000000000000000000000800 00000000000000000000000000000000A000010 000000000000
25		Mask 26: 000000000000000000000000FFFF 00000000000000000000000000000000FFFFF0 000000000000
30		Value 26: 0000000000000000000000000800 00000000000000000000000000000000A0000E0 000000000000
35		Mask 27: 000000000000000000000000FFFF 00000000000000000000000000000000FFFFF0 000000000000
		Value 27: 0000000000000000000000000800 00000000000000000000000000000000A0000F0 000000000000

For each subscriber, the rate governors can be defined according to the service they receive. In this example, premium subscribers can burst to 150% of their normal rate, while other subscribers are limited to the normal rate.

- Peak Rate "Premium": 30 Mbit/s  
Profile Rate "Premium": 20 Mbit/s  
Burst Limit "Premium": 30 Mbit  
Peak Rate "Standard": 10 Mbit/s  
5 Profile Rate "Standard": 10 Mbit/s  
Burst Limit "Standard": 15 Mbit  
Peak Rate "Value": 1 Mbit/s  
Profile Rate "Value": 1 Mbit/s  
Burst Limit "Value": 1.5 Mbit  
10 Premium subscribers can have their traffic assigned to the third highest priority queue or third party output buffer 570, while standard and value subscribers can be assigned to the lowest priority or fourth priority output buffer 572.

#### BACKBONE NETWORK INTEGRATION

- 15 Quality of service (QoS) is most powerful when it can be managed globally across an entire network, and the present invention provides unparalleled opportunities for global QoS management across an entire backbone network. The basis for this integration is IP's differentiated services (diffserv) architecture.

#### 20 Application Support for Diffserv

- SOIs 140 can support two applications that can significantly benefit from quality of service support: voice over IP and T1/E1 over IP. In both cases, the service provider can configure the application to mark its packets with a particular diffserv code point. These settings allow either application to take advantage of expedited forwarding, assured  
25 forwarding, or class selector prioritization throughout the IP network with the present invention. In addition, the SOI's VoIP implementation supports the setting of DSCP values on a call-by-call basis on command of the media gateway controller. This feature allows, for example, giving special priority to specific calls (e.g. E911 service).

## Creating Service Level Agreements

The Transceiver Node (TN) 120 provides extensive support for managing service level agreements (SLAs) with subscribers. Although the TN 120 is necessarily only one component in an overall agreement, as the access network, it is critical. The following  
5 examines how the TN contributes to SLAs and how the above teaching can support SLAs through its so-called *quality of service* (QoS) and management functionality.

## Components of an SLA

Service level agreements are typically more common with private network  
10 technologies such as ATM or Frame Relay. The power and flexibility of the TN's 120 QoS management, however, permits those same concepts to be extended to IP access networks. The same components that are part of traditional ATM or Frame Relay SLAs can be part of an TN-managed SLA.

### 15 Definitions used herein:

- Peak Rate. The maximum rate at which the network will accept traffic bursts from the user, expressed in bits per second. The network discards traffic that exceeds the peak rate.
- Sustained Rate. The minimum throughput that the network will provide to the  
20 user, expressed in bits per second.
- Burst Size. The amount of traffic that the network will accept without pause at the user's peak rate, expressed in bits.
- Maximum Latency. The worst-case delay the user's traffic will experience as it traverses the network.
- 25 • Loss Rate. The percentage of traffic conforming to the peak rate, sustained rate, and burst size that the network may discard.

Of course, service providers can include other elements in their service level agreements. The Transceiver Node 120 provides a wealth of features that a service

provider may position as value-added services. The TN 120 supports services such as the following:

- 5       • Application Prioritization. Giving priority to key network applications (e.g. Virtual Private Network traffic).
- Enhanced Statistics. Providing detailed traffic profiles and statistics to assist the user in network growth planning.
- Active Monitoring. Continuously monitoring user traffic to provide early detection of network application faults (e.g. Web server failures).
- 10      • Network Security. Providing encryption of traffic to the subscriber.

15       This part of the description focuses on traditional SLA performance metrics. It examines how the Laser Transceiver Node 120 contributes to network performance, and how to provision downstream QoS management to meet SLA requirements. The table below lists key parameters and values used in equations throughout this part of the description.

#### Inherent Link Characteristics

C	Link Capacity (500 Mbit/s)
$\tau$	Superframe Period (8 ms)

#### 20   Rigorous SLAs and Oversubscription

      Because business requirements differ among service providers and among subscribers, the Transceiver Node 120 allows providers significant flexibility in enforcing SLA performance metrics. Some deployments can require ironclad service level agreements; those environments require a conservative provisioning strategy.

25   Conservative provisioning can provide extremely tight performance guarantees, but it generally results in a lower overall network utilization and, ultimately, greater capital expenditures.



In other deployments (residential Internet access, for example) SLAs are not common and may not be desirable. In those environments a more aggressive provisioning strategy may be effective. In general, meaningful SLAs are usually not enforceable when a network is provisioned aggressively; the resulting networks, however, may be operated at much higher utilization.

This part of the description considers both strict SLAs and slightly relaxed SLAs. Relaxed SLAs allow a modest amount of oversubscription of network resources; in exchange, the service provider cannot offer rigorous guarantees for all aspects of network performance. Oversubscription typically means that the service provider has promised somewhat more bandwidth than he has the technical capacity to deliver. Since most users typically do not continuously utilize all of their promised or guaranteed bandwidth, the unused portion of the guaranteed may be temporarily assigned to other users.

#### Downstream Performance

The flexibility of the Transceiver Node 120 provides extensive flexibility in controlling downstream performance, and there are many different ways to provision downstream links. This section considers a typical configuration for environments in which service level agreements are more common – Internet access for businesses. To focus on the key parameters, this discussion makes several simplifying (but not unrealistic) assumptions.

- Internet data traffic is classified separately from other applications. Separate classifiers are used for specific applications such as voice or T1/E1 over IP.
- Each subscriber's data traffic is classified and policed independently. This assumption requires that one classifier and one policer be dedicated to each of the 16 subscribers on a channel.
- All constant bit rate (CBR) traffic (e.g., voice on IP, T1/E1) is policed by a sustained rate and burst size only; peak rates are not used for this traffic. (Policers for non-data traffic have their WRED parameters for out-of-profile

traffic set to discard all out-of-profile packets; setting both the minimum and maximum thresholds to zero accomplishes this action.)

- Data traffic that is not time critical (web surfing, file downloading, etc.) is assigned to the lowest priority output buffer.
- 5      • All 16 subscribers' data traffic policers have the same WRED parameters for in-profile traffic, and differ for out-of-profile traffic only in the maximum discard probability.

Recommended values for the WRED parameters include the following (see  
10      Figures 9 and 10 for definitions):

- In-profile minimum threshold,  $Th_{min}$ , of 50000 bytes
- In-profile maximum threshold,  $Th_{max}$ , of 150000 bytes
- In-profile maximum drop probability  $P_{max}$  of 26 (corresponding to a probability of 25/256)
- 15      • Out-of-profile minimum threshold,  $Th_{min|out}$ , of 10000 bytes.
- Out-of-profile maximum threshold,  $Th_{max|out}$ , of 30000 bytes

With these assumptions, the following parameters can characterize downstream performance.

#### Downstream Channel Characteristics

$C_D$	Downstream link capacity; the physical link capacity less sustained rates for all constant bit rate traffic
$H_D$	Sum of the burst sizes for all non-data policers

20

#### Downstream Configuration parameters (per Subscriber)

$B_D$	Downstream Burst Size (bit)
$P_D$	Downstream Peak Rate (bit/s)
$R_D$	Downstream Sustained Rate (bit/s)
$W_D$	Downstream Maximum Discard Probability (unit-less)

Both strict SLAs and lenient SLAs are possible. Strict SLAs require configuration that satisfies the following constraint.

- The sum of the peak rates for all subscribers must be less than the link capacity.  $[\sum P_D < C_D]$

With that constraint, SLA parameters are easy to derive from configuration values.

SLA Metric	TN Configuration Parameters
Peak Transmission Rate	equal to Downstream Peak Rate $[= P_D]$
Sustained Transmission Rate	equal to Downstream Sustained Rate $[= R_D]$
Transmission Burst Size	equal to Downstream Burst Size $[= B_D]$
TN Downstream Latency	no more than the time spent waiting for non-data traffic plus the time to transmit the out-of-profile maximum threshold worth of data $[= H_D/C + Th_{max/out}/C_D]$
TN Downstream Loss Rate	0

5

Lenient SLAs require a less strict configuration constraint, namely the following.

- The sum of the sustained rates for all subscribers must be less than the link capacity.  $[\sum R_D < C_D]$

In the lenient case, closed form equations for SLA parameters are not possible.

10 The following rules provide approximate bounds for those parameters.

SLA Metric	TN Configuration Parameters
Peak Transmission Rate	either the Downstream Peak Rate or at least the weighted share of excess link capacity (capacity above the Downstream Sustained Rates of all TNs, whichever is smaller $[\geq \min(P_D, R_D + (C_D - \sum R_D) * B_D * W_D / \sum (B_D * W_D))]$
Sustained Transmission Rate	equal to Downstream Sustained Rate $[= R_D]$

15

Transmission Burst Size	equal to Downstream Burst Size $[= B_D]$
TN Downstream Latency	no more than the time spent waiting for non-data traffic plus the time to transmit the out-of-profile maximum threshold worth of data $[= H_D/C + Th_{max/out}/C_D]$
TN Downstream Loss Rate	0

It should be understood that the foregoing relates only to illustrate the embodiments of the present invention, and that numerous changes may be made therein without departing from the scope and spirit of the invention as defined by the following

5 claims.